JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

E-commerce recommendation system using reverse image search and questionnaire

Gayatri Chaure[1], Aishwarya Dani[2], Vaishnavi Hapse[3], Pradnya Jadhav[4], Sharada Ohatkar5

[1,2,3,4,5]MKSSS's Cummins College of Engineering for Women, Pune

***Abstract*—** The potential uses of a technology are vast and continually expanding, as our knowledge and comprehension of it increases over time. An interesting and promising area that is currently being explored is the reverse image search. This technology has a wide range of applications that can greatly simplify our lives in various business sectors, as well as in our everyday routines. The way it works is by accepting images as the search input, rather than text, and returning similar images as the output, thereby making it easier for users to find what they are searching for. In this paper, the VGG16 model, which includes a combination of Convolutional Neural Networks (CNNs) and k-nearest neighbors (KNN) algorithm, is utilized to conduct reverse image searches. The results of the reverse image search are presented, and a short questionnaire is provided to the user to obtain feedback. The questionnaire inquiries about the product category, and the information gathered from the users' responses is used to filter the relevant results by category.

Keyword: Convolutional Neural Networks, k-nearest neighbors (KNN), reverse image search

I. INTRODUCTION

In the world of online shopping, there is an ease in searching for any product from any corner of the world. We live in a digital world where everything is openly accessible to anyone. Online shopping is now part of our life . It solves various problems and difficulties which were there earlier. Yet, all these things are readily available on e-commerce websites or marketplaces to obtain a certain product, one must browse the entire list of items offered on the online buying platform. There are various methods for searching for products on an online purchasing platform, including text input, speech input, and image input. The first technique being used is text-based search, we called it textual search. In textual search if we want to search for any clothing category, one has to use a keyword or specific phrase to describe the image. then it will compare that with the available database. The difficulty in this technique was if user input is not matched with the available dataset, then non relevant images were shown as output.

The technology that came next is speech input searching. With the help of speech, the user can give input to the system about what he/she wants to buy and according to that the output appears. The difficulty in this technology is that if input contains background noise or the voice of the user is not clear or mispronounced then there is a chance of non-relevant output. The third technique which is widely used nowadays is image search. Consider an example from the fashion industry, if one has to search tops, sometimes non-relevant images come as output. By taking a photo of

the item you're looking for, uploading it to an online buying site, it will give you much more relevant output. To get more accurate results and to give better experience for the user, we combine image search with the questionnaire. Choosing the right name from a drop-down menu or questionnaire would help to get more accurate results.

For advertising purposes and to give their customers the most thorough view of their products, online shops largely rely on photographs. It gets harder and more expensive to manually annotate every new image that is created as a result of the exponential growth in the number of photos that are already stored in databases and servers. We mix picture search with the questionnaire to produce more precise results and a better user experience. Choosing the right name from a drop-down menu or questionnaire would help to get more accurate results.

II. RELATED WORK

1. The article [1] describes the creation and application of a Convolutional Neural Network (CNN) for an image search engine in the fashion industry. The algorithm uses both eu- clidean and cosine distance classification methods. the KNN classifier is able to classify and match interest points more accurately than simply using the extracted interest points and learned feature vectors. As a result, less processing power is required and the resulting accuracy achieved is approximately 96.08.
2. The paper [2] details a technique that utilizes SIFT descriptors to match a reference image. The method involves employing a KNN classifier with a pre-trained feature vector to classify and match the interest points that were extracted. By transforming the image into a small feature dimension, the method effectively extract relevant features and reduces the problems size. As a result, the accuracy achieved is approximately 79
3. The article [3] examines three models, namely VGG-19, INCEPTION V3, and ResNet-50, and reports their corre- sponding accuracies of 76.1
4. The study presented in [4] employed a CNN and SVM model, resulting in a 90
5. The study of [5] found that the Pre-Trained CNN Model method could handle orientation information, but its com- putational complexity is 28.2 times longer than the p-Hash method when producing a feature vector. Despite this, CNN is more effective than p-Hash, with an average of 1.4 times better precision, recall, and F-measure.
6. In the paper[6] they conclude that the accuracy and precision of an image search system largely depend on the accuracy of the image features obtained.
7. The article [7] covers the fundamental principles of neural network operations in image processing. The study reports that a model with 32 hidden layers achieved an accuracy of approximately 82.20

III. ARCHITECTURE

Andrew Zisserman and Karen Simonyan introduced the VGG model in 2013 as part of the Visual Geometry Group (VGG) at Oxford. They created a prototype for the 2014 ImageNet Challenge, and the VGG model differed from previous models in a few key ways. Firstly, it utilized a small 3x3 receptive field with a 1-pixel stride, in contrast to the larger 11x11 receptive field with a 4-pixel stride used in AlexNet. Despite its smaller size, the 3x3 filters in VGG effectively captured features of a larger receptive field.

VGG16 model is used for feature extraction and the KNN algorithm for classification. Model requires images of dimensions 224x244 with RGB channels . The VGG16 model processes the input image and generates a vector consisting of 1000 values as output.The initial layer employs 64 convolutional filters sized 3x3 with identical padding. The second layer is also a convolutional layer with 64 filters of the same size and padding. Following this, a max pooling layer with a stride of (2,2) is applied to reduce spatial dimensions while retaining important features. The network then utilizes two convolutional layers with 128 filters of size 3x3, followed by another max pooling layer with the same stride. This is succeeded by two convolutional layers with 256 filters of size 3x3, and then two sets of 3 convolutional layers with 512 filters each, and a max pooling layer.Max pooling is a form of pooling layer utilized in VGG16, which involves downsampling feature maps by selecting the maximum value within a local region. This process helps to reduce the spatial dimensions of the feature maps, promotes feature selection by retaining the most prominent features, and enhances translation invariance and robustness to small input variations These multiple sets of convolutional layers enable the network to capture more complex features from the input image . VGG16 incorporates the Rectified Linear Unit (ReLU) activation function after each convolutional and fully connected layer, excluding the output layer. ReLU introduces non-linearity into the network, allowing it to learn complex features and mitigate the vanishing gradient prob- lem. ReLU is widely utilized in deep learning architectures due to its effectiveness in improving model performance by enabling better representation of non-linear relationships in data.

Due to its ease of use and high degree of classification accuracy KNN is used for the classification operation, it has found widespread application in numerous fields. The KNN algorithm’s main principle is to first calculate the distance (Euclidean distance). The term ”distance” refers to the length of a line segment between two points. The Euclidean Norm calculates a vector’s magnitude. The equations are the same, and the magnitude of a vector is essentially its length. Find

Fig. 1. VGG16 model architecture.

the K neighbours with the distance or similarity that is closest to the distance between the query sample and the training sample, or similarity, and the query sample.

IV. OUR APPROACH

Reverse image search is a revolutionary technology that is transforming the landscape of e-commerce recommendation systems. Through the utilization of machine learning and computer vision, reverse image search enables online retail- ers to analyze and identify images of products, whether up- loaded by users or sourced from the internet. This advanced approach allows for more accurate and relevant product recommendations to customers, eliminating the need for manual keyword searches. By simply uploading or click- ing on an image, users can find visually similar products, resulting in an improved shopping experience and higher chances of conversions. In this article, we will explore the advantages and applications of reverse image search in e-commerce recommendation systems, as well as delve into the cutting-edge technologies and techniques that underpin its functionality.

The input is passed through a stack of two convolutional layers with 3x3 filters, with 1-pixel padding added after each convolutional layer to maintain spatial dimensions. The result is converted into a fully connected layer after softmax operation is done on it. the final output is in the form of array which is used for comparison purpose.Overall, this convolutional neural network (CNN) architecture is well- suited for feature extraction tasks. The use of smaller filter sizes may allow the network to capture finer details in the input images.Before using the model we are processing the product images. We are storing the adresses of features of the images and product images in the CSV file. User will upload the query image. On this image we are doing the feature extraction and converting it into feature vector which is further used for the comparison purpose.

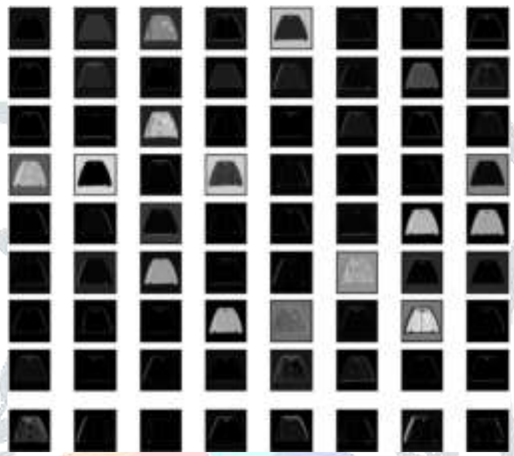

Fig. 2. Feature extraction of product image.

In the project, the dataset is stored in a CSV file, which is a tabular format that can be easily managed and processed. The images corresponding to the dataset are stored in a separate folder on the device. The paths of these images in the folder are stored in the CSV file, along with additional data such as image characteristics like color, gender, and type of outfit. This allows for a systematic organization of the dataset, making it easier to retrieve and manipulate the data.The main component of the project is a ”Questionnaire” that is provided to the user. This questionnaire contains questions about the image the user wants to search for. The user provides responses to these questions, which help in narrowing down the search criteria and refining the search results. For example, if the user is looking for a red dress for women, the questionnaire may ask questions related to color, gender, and type of outfit. The responses from the questionnaire act as filters that are applied to the dataset in the CSV file.

By applying the filters to the data in the CSV file, the dataset is reduced to a smaller subset that matches the criteria provided by the user in the questionnaire. This reduction

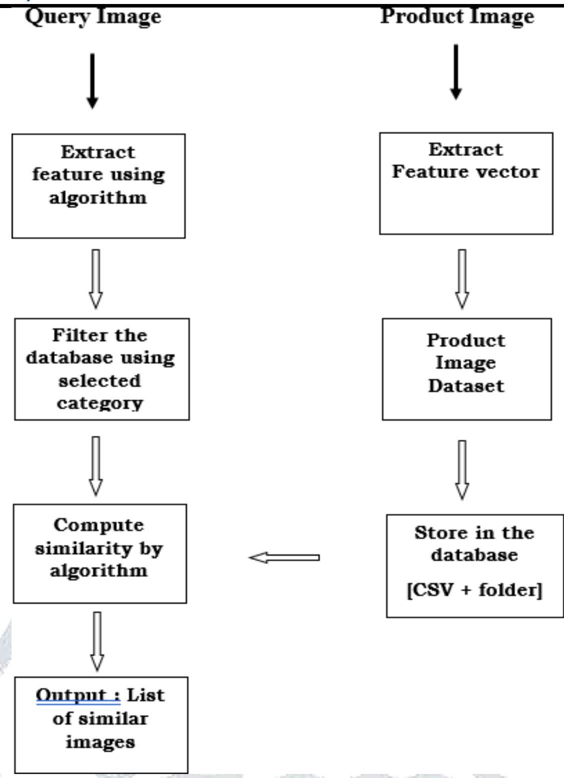

Fig. 3. Block diagram – reverse image recommendation system.

of the dataset helps in saving time and resources, as the search operations are performed only on the filtered data, eliminating the need to search through the entire dataset. Moreover, it increases the accuracy of the search results, as the filtered data is more relevant to the user's requirements. Once the dataset is filtered, the reverse image search operations are performed only on the filtered data. This optimizes the search process, as the search is now focused only on the relevant subset of the dataset. The filtered data can be used as a reference for further analysis, retrieval, or visualization of the images that match the user's criteria. Overall, the use of a questionnaire and CSV file for filtering the dataset helps in improving the efficiency and accuracy of

the reverse image search process in the project.

The k-nearest neighbour samples that are closest to the ob- ject d to be classified are found by searching the training set

S. Distance metrics are used, such as Euclidean distance, in "Nearest Neighbours". The Euclidean distance of two objects $X_1 = (x_{11}, x_{12}, . . . , x_{1n}) and X_2 = (x_{21}, x_{22}, . . . , x_{2n})$ is calculated by formula shown in below equation (1)

$$\text{dist}(X_1, X_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2}. \quad .....(1)$$

The choice of distance measurement is crucial, and it is possible to measure other types of distances as well. The distance is compared. The greater the similarity, the smaller the distance. We can sort the list by distance and then take the first k values from it. As it will display the first k images from the dataset that are closest to the query image.Using NumPy is going to be quite a bit faster. For this NumPy library can be used. This library is used for manipulating multidimensional arrays in a very efficient way. np.linalg.norm can be used for distance measurement between numpy arrays. By de- fault, numpy.linalg.norm() function computes the euclidean distance that is square root of the sum of the squared vector values Therefore, in order to compute the Euclidean Distance we can simply pass the difference of the two NumPy arrays to this function The below fig.5 shows how to implement the Euclidean distance function using code. The euclidean distance can be stored in a list, sorted in ascending order and Images of first k values in the sorted list can be given as result.

```
pseudo code :
def euclidean_distance(numpy_array_image1, numpy_array_image2):
  result=np.linalg.norm(numpy_array_image1 - numpy_array_image2)
  #numpy_array_image1 and numpy_array_image2 are feature vectors of two images
return result
```

Fig. 4. Pseudo code for euclidean distance calculation

V. RESULTS

The below fig 4 shows the user interface where the user can upload the image that he/she wants to search . The query image will be displayed on the website and below that the resultant images will be displayed on the interface. The user interface has a questionnaire option before actually submitting the query image.

Fig. 5. Web interface for query submission

The way in which our application is implemented is depicted below. We chose 10 classes from our dataset, some of which had relatively more sample training photos than others. In order to test the software, we used an image called ”jeans.jpg” that wasn’t included in the test/train dataset, as shown in above figure. This picture, along with every other picture used in the application testing procedure, was chosen at random from the internet. It serves as an illustration of the type of image a user might enter into the system when submitting a query. Our system has ranked the photos based on which ones are the most similar, and those shirts are what

are shown below in figure. Each image in the results consists of euclidean distance and it is shown just below every image.

Fig. 6. Result of system

VI. CONCLUSION

We developed and trained our own Convolutional Neural Network (CNN) using a custom dataset. Once we achieved a relatively high level of accuracy, we implemented a reverse search engine using Euclidean distances. To extract features, we utilized VGG16, and for classification, we employed the KNN algorithm. On the user interface, users can upload a query image to search, and the resulting images are displayed beneath the query image. Additionally, we included a ques- tionnaire option for users before submitting their query im- age. Our observations indicated that categories with a higher number of training images per class generally had a higher degree of accuracy during experimentation, suggesting that having an ample amount of training images per class is crucial to prevent the CNN model from overfitting.

VII. FURTHER WORK AND IMPROVEMENTS

In current web application, our dataset consists of product images havingfile size of 2kb to 3kb. It can be improved further to get better resolution of the images in the result. The robustness of the model can also be increased by using a dataset with more images. Other distance measures, such as hamming distance, cosine distance, etc., may be incorporated to complete the similarity search component after the CNN has been trained satisfactorily. Reverse image search can be integrated into mobile e-commerce apps, allowing users to capture and upload images of products they find interesting or want to purchase directly from their mobile devices. This can enhance the convenience and usability of reverse image search for mobile shoppers, who may prefer visual search over typing keywords on a small mobile screen.

REFERENCES

[1] K. F. Mawoneke, X. Luo, Y. Shi and K. Kita, ”Reverse Image Search for the Fashion Industry Using Convolutional Neural Networks,” 2020 *IEEE 5th International Conference on Signal and Image Processing (ICSIP)* , Nanjing, China, 2020, pp. 483-489

[2] S. R. L. Suganthi and Hanumanthappa M., ”Image mining using SIFT based object identification and tagging,” *2016 International Conference on Computation System and Information Technology for*

Sustainable Solutions (CSITSS), Bengaluru, India, 2016, pp. 261-266

[3] Fengzi Li, Shashi Kant, Shunichi Araki, Sumer Bangera, Swapna Samir Shukla, “Neural Networks for Fashion Image Classification and Visual Search”, BT5153 Applied Machine Learning Term Paper, *MSBA, NUS, 2020*

[4] Arif Rahman a,b, Edi Winarko a, Khabib Mustofa , “Product image retrieval using category-aware Siamese convolutional neural network feature”, *Journal of King Saud University – Computer and Information Sciences 34(2)*, April 2022

[5] Hanaa Al-Lohibi, Tahani Alkhamisi, Maha Assagran, Amal Aljo- hani, and Asia Aljahdali, “A Reverse-Image-Search Application”, *Advancesin Distributed Computing and Artificial Intelligence Journal Regular Issue Vol 3 n.3 2015*

[6] B. Maleki Shoja and N. Tabrizi, ”Customer Reviews Analysis With Deep Neural Networks for E-Commerce Recommender Systems,” in *IEEE Access, vol. 7, pp. 119121-119130*, 2019

[7] Google image search: http://images.google.com/.